LBNL-44045

Review of Sierra Research Report "Development of a Proposed Procedure for Determining the Equivalency of Alternative Inspection and Maintenance Programs"

Tom Wenzel and Robert Sawyer
Lawrence Berkeley National Laboratory
September 23, 1998

Summary of I/M Evaluation Methodology Proposed by Sierra Research

The Sierra report proposes a methodology for states to use to demonstrate that their Inspection and Maintenance (I/M) programs are obtaining emissions reductions similar to those estimated by the MOBILE emissions factor model. The proposed methodology involves comparing fleetwide test average emissions of a state I/M program with those of a benchmark program (Sierra suggests using the Arizona IM240 program for the benchmark program). The fleet emissions are determined based on the final (passing) IM240 test results of a stratified random sample of 1,600 vehicles. Average fleet emissions are determined based on the age distribution of the fleet (Sierra suggests using three age groups, pre-1975, 1975-1980, and post-1980). The fleet emissions are then compared with the fleet emissions of the benchmark program, reweighted using the evaluated program's vehicle age distribution.

Programs that use an alternative I/M test (e.g. ASM, idle, etc.) and that cannot test a random sample of IM240s, must first determine the correlation between their test and the IM240, based on a stratified random sample of 700 to 800 vehicles measured on both tests (depending on the type of alternative test used). Then fleet emissions are determined based on the final (passing) IM240 test results of a stratified random sample of 3,700 to 8,000 vehicles (again, depending on the type of alternative test used). The weighted average emissions of this fleet are then compared with those of the reweighted benchmark program.

Four recruitment options are proposed to ensure that the sample of vehicles tested is random and not biased. Evaporative HC emissions are compared by converting functional pressure test failure rates of both the evaluated and benchmark program to evaporative emissions rates, using the MOBILE model. To evaluate the effectiveness of visual and functional inspections, use of at least 30 undercover or covert audit vehicles is recommended for each visual and functional inspection.

Our review of the Sierra report is divided into three sections, based on the type of comments: comments on details of the proposed methodology; general comments on the overall methodology; and a summary of how one would conduct a more accurate evaluation of a state's I/M program.

Details of the Proposal

1. *Data used.* Sierra used the test results from the CARB I/M Pilot Program in its determination of the minimum number of vehicles needed to demonstrate correlation and to obtain a random sample. This program consisted of FTP, IM240, ASM, and idle testing of the same fleet of in-use vehicles. A great effort was made to ensure that all vehicles contacted participated in the study. About 60% of the contacted vehicles actually participated; this rate is much higher than that of CARB's on-going in-use surveillance testing (which typically tests only 10 to 15% of the vehicles solicited). However,

the potential still exists for selection bias in the sample (that is, owners of cars suspected of being high emitters may have refused to participate at a higher rate than clean vehicles).

2. *Bias in concentration-based measurements.* Analysis of remote sensing and IM240 average emissions by vehicle model indicate that concentration-based measurements (idle, ASM, remote sensing) introduce a bias against vehicles with higher fuel economy/smaller engines (Wenzel 1998a). That is, smaller, fuel-efficient vehicles can afford to emit higher concentrations of pollutants out their tailpipe while still meeting mass-based (gram per mile) emissions standards. This complicates any attempt to correlate concentration-based alternative tests to the IM240. Sierra did examine the effect of correcting concentration-based emissions measurements by vehicle weight, in their calculations of minimum sample sizes required; however, since this correction had little effect on their results, they ignore the effect (p. A-23). We recommend that this correction be made when comparing concentration-based measurements to mass-emissions based measurements.

3. *Stratified sampling.* Sierra suggests stratifying vehicles into three MY groups, to account for changes in catalyst technology (and increasingly stringent new car certification standards) over time. However, Sierra includes all three-way catalyst equipped vehicles in the same MY group (1980 and newer). This grouping ignores technological advances, such as the development of fuel injection, that may not have had much impact on new car emissions, but likely had a significant effect on emissions control durability and in-use emissions. We suggest that the MY80+ vehicle group be divided, at least into a MY80-88 group (mainly carbureted) and MY89+ group (fuel-injected), and perhaps a third (MY93+, port fuel-injected) and fourth (MY96+, Tier 1 OBD2 equipped). Alternatively, vehicles can be grouped by fuel system (carbureted vs. throttle body fuel injection vs. port fuel injection) and emissions standards (Tier 0 v. Tier 1), independent of MY, if possible. Also, since high emitters account for a large portion of pre-test fleet emissions, and the purpose of the I/M program is to reduce emissions from high emitters, perhaps the sample should be stratified on pre-test emissions (initial pass vs. initial fail) as well. Otherwise one runs the risk of testing a "random" fleet that has very few high emitters.

4. *Comparison of older model years.* Most states are not going to require IM240s of older vehicles (1980 and older); rather, idle testing will be required of older vehicles (as in Arizona and Colorado). Emissions from older vehicles, as measured by idle tests, cannot be easily incorporated into fleet average emissions as measured by IM240 testing. Sierra recommends using correlation equations to convert idle emissions to IM240 emissions (p. 27). Rather than attempting to convert idle emissions to IM240 emissions, we suggest that, in an interstate comparison, older vehicles tested under idle be compared independently of newer vehicles tested under enhanced tests. (Our alternative method to evaluate I/M programs, discussed in Section 3, would apply the same test methodology to all vehicles regardless of age).

5. *Recruitment of random sample.* Sierra proposes four possible methods of recruiting vehicles for the random sample: 1) at a test-only inspection station, at the time vehicles pass their I/M test; 2) called in for additional testing, with non-participants denied registration renewal; 3) unscheduled mandatory roadside testing; and 4) voluntary participation, as long as the distribution of emissions-related defects in the sample matches those in the fleet, as determined by mandatory roadside inspection (p. 22). The first of these appears to be the best method of recruiting vehicles; the other options may not be practical. For example, Arizona required that vehicles identified by remote sensing as being suspected high emitters come in for an unscheduled I/M test. The penalty for non-compliance was suspension of vehicle registration. About half of the vehicles identified as high emitters by remote sensing did not report for additional I/M testing (personal communication, John Brown 6/97). The threat of registration suspension was not enough to get these motorists to comply with the program (although the Arizona Motor Vehicle Department had the incorrect addresses for about one-third of the motorists that did not respond (personal communication, Dan Grubbe 9/98). Arizona found it politically difficult to enforce strictly the registration suspension (in part because the false failure rate, based on remote sensing measurements, was so high). Another example is

California's experience in conducting roadside testing. For a study conducted in 1994, only 30% of motorists pulled over allowed inspectors to perform a voluntary underhood visual inspection. Remote sensing measurements of these vehicles indicated that their CO and HC emissions were more than twice as high as the vehicles that agreed to the roadside visual inspection (Stedman et al 1994). California has determined that it cannot legally require all vehicles selected to undergo roadside testing, so it is likely that there will continue to be a recruitment bias in roadside emissions testing.

When selecting a random sample of vehicles, care must be taken to account for any seasonal variation in emissions. Analysis of average daily IM240 emissions indicates that average CO emissions tend to be higher in warmer summer months, and lower in colder winter months; HC emissions show a similar but smaller trend, while NOx emissions show the opposite trend (Wenzel et al 1998). It is thought that a combination of temperature and oxygenated fuel accounts for the seasonal variation in IM240 emissions; another possibility is the temperature and humidity correction factors for NOx emissions (Joy 1998). In addition, the types of vehicles brought in for I/M testing at different times of the year may affect sampling procedures. For example, it has been reported that all government fleet vehicles in Ohio are tested in the same few months (personal communication, Ed Glover 9/98). This affects the overall vehicle age distribution, and perhaps the degree of maintenance performed on the average vehicle, in these months (government fleet vehicles are thought to be newer, and perhaps better maintained, than the average vehicle). Care should be taken in the sampling of vehicles from the evaluated program to reduce the impact of any of these factors in affecting vehicle emissions relative to those in the benchmark program.

Selection bias is difficult to eliminate if testing is not mandatory. However, one can attempt to determine the direction and degree of bias by soliciting more vehicles than necessary, and then comparing subsequent emissions measurements (through the regular I/M program) of the participating vehicles with those of the non-participating vehicles (personal communication, Robert Slott 9/98).

6. *Distribution of in-use emissions.* Sierra demonstrates that IM240 emissions of in-use vehicles, as measured in the CARB pilot project, are not normally distributed, and more closely match a lognormal distribution than a normal or gamma distribution. Our analyses of multiple sets of in-use data confirm this finding. Emissions for HC, CO, and NOx, each follow the lognormal distribution, although the distribution for each pollutant has a characteristic shape. In addition, the distributions from different datasets (AAMA FTP Bag 3, CARB LDVSP 13 FTP Bag 3, EPA/ATL IM240, Arizona IM240, Colorado IM240) are somewhat different. (Differences in the emissions distributions under each testing program can be attributed to a number of factors including vehicle preconditioning, test cycles used, recruitment bias, and vehicle age. We attempt to account for vehicle age by examining vehicles with roughly similar mileage). All of these factors can affect the evaluated program's emissions relative to those of the benchmark program.

7. *Correlation between IM240 and alternative test.* Sierra includes several figures demonstrating the relationship between FTP and IM240 (R^2 of 0.89/0.66/0.78 for HC/CO/NO_x), ASM2525 (0.77/0.72/0.53), and idle (0.64/0.26/-- for low speed and 0.59/0.66/--) for high speed) emissions testing on the same vehicles. (It would be interesting to see another set of figures, comparing the emissions reductions for individual vehicles as measured by the IM240 and alternative tests). The relatively low correlations are due to two effects: 1) intrinsic variability of vehicle emissions, due to intermittent emissions controls failures, and 2) differences in the vehicle operating conditions, and therefore emissions, measured by each test. Vehicle emissions can be quite variable even in replicate FTP tests in carefully controlled conditions; high-emitting vehicles in particular show high emissions variability in replicate testing (Knepper et al 1992).

The correlations Sierra presents are based on carefully controlled I/M tests with consistent vehicle preconditioning. One would expect a larger number of false passes from field tests conducted in operating I/M lanes, where preconditioning and other test procedures may not be as carefully controlled (for example, see Heirigs and Gordon 1996). Therefore the agreement between FTP tests and I/M tests conducted in the field would not likely be as good as observed between FTP tests and controlled I/M tests.

Sierra recommends that a certain number of vehicles be tested under both IM240 and an alternative test to demonstrate correlation. Presumably the correlation equation would then be applied to the alternative test values to obtain IM240-equivalent emissions. However, Sierra sets no criteria for whether the two tests are correlated; is an R^2 of 0.66 (Figure 2) sufficient to demonstrate correlation? Or is the R^2 value irrelevant, so long as an equation to convert alternative test values to IM240-equivalent values is derived? Sierra should explain what constitutes demonstrating correlation, and what should be done if correlation is not adequately demonstrated.

As discussed above, vehicle emissions exhibit a large degree of test-to-test variability. Perhaps a better indicator of equivalency of an alternative test to the IM240 is a measure that focuses not on the emissions of individual vehicles, but the degree to which both tests identify high emitters. Sierra mentions such a measure in a footnote (p. 7): a combination of false failures and excess emissions measured by the alternative test, relative to the IM240, could be used.[1] We suggest that EPA consider recommending to states other means of demonstrating alternative test equivalency, rather than correlations.

Demonstration of equivalency between IM240 and alternative emissions tests does not mean that an I/M program using an alternative emissions test will achieve the same emissions reduction as an IM240 program. For example, several simple engine adjustments can be made to make a vehicle pass an idle test, whereas more detailed repairs may be required to make the same vehicle pass an IM240. Sierra makes this point on p. 15.

Sierra recommends developing correlations of alternative tests to the IM240 in order to obtain an equation to convert emissions as measured under the alternative test to IM240-equivalent emissions. This step needs to be done to follow Sierra's methodology of comparing all I/M programs to a benchmark IM240 program. However, allowing states to evaluate their programs using emissions as measured under their own program, as we recommend below, removes the need to convert test results to IM240-equivalents, and the need to develop correlation equations.

8. *Using remote sensing data.* Sierra demonstrates that the difference in average emissions measured by remote sensing instruments at two different sites is comparable to the difference in average emissions between old and new vehicles at the same site (p. 17). Consequently, Sierra recommends that remote sensing data not be used to predict FTP or IM240 emissions, unless site-specific correlations are developed. Two different contractors in the Sacramento Pilot Project, using two slightly different instrument technologies, generated the remote sensing data Sierra uses. Several concerns about the accuracy of many of the instruments used in this study have been raised. These data are perhaps not the best remote sensing data to demonstrate the point.

A certain amount of variability in remote sensing average emissions by MY and site is to be expected. The variability can be due to three factors: instrument error, differences in vehicle operating mode, and differences in the vehicle populations driving by those sites. It is our opinion that instrument error can be greatly minimized, if not eliminated, if reasonable care is taken in setting up and calibrating the instruments. This may or may not have been the case in the

1. In the same footnote Sierra states that "the objective of the I/M evaluation process is to determine FTP emissions". We disagree with this statement; we believe that the I/M evaluation should examine reduction in in-use, rather than certification test, emissions. FTP emissions include cold starts, optimized test fuels, controlled temperatures, trained drivers, etc.; none of these testing protocols are included in EPA's Enhanced I/M Testing Guidelines.

Sacramento Pilot Project. In terms of differences in vehicle operating mode, Sierra explains that the sites selected for their comparison were determined to be "A" sites by the California BAR, so dramatic differences in vehicle operating mode are unlikely. On the other hand, the vehicle populations driving by each of the remote sensing sites may vary in terms of distribution of vehicle models, mileage accumulation, or owner maintenance practices. It is not clear how much of the site variability noted by Sierra is due to differences in the vehicle populations measured at each site. We have shown that IM240 emissions measurements are subject to "site variability" as well; average emissions by vehicle model are consistently higher at an Arizona IM240 test station located in a relatively low-income area than those at a station located in a relatively high-income area (Wenzel and Ross 1997). These differences are likely due to differing vehicle conditions and maintenance practices in the two sites.

Overall Methodology

Sierra recommends comparing the emissions of a fleet after it has completed its I/M program to those of a benchmark program, rather than measuring the actual emissions reductions (pre- and post-repair) within a program. This is recommended in part because it is very difficult to determine baseline emissions from a non-I/M case, from which to calculate emissions reductions. However, comparing average emissions across I/M programs in different states involves other problems; many factors that affect emissions can raise or lower a program's average emissions relative to those of the benchmark program:

- I/M equipment and testing procedures (cutpoints used, preconditioning procedures, use of fast-pass/fast-fail algorithms, frequency of calibration, etc.);
- fleet composition (vehicle type, model, age, and mileage);
- fuel composition (reformulated gasoline, winter oxygenates);
- driving conditions (frequent driving under heavy load, as in mountainous areas);
- environmental conditions (ambient temperature and humidity, altitude); and
- socioeconomic factors (which can affect average vehicle condition and maintenance practices).

Sierra recommends using the MOBILE model to correct for some of these differences (fleet composition, fuel composition, and fraction of fleet subject to I/M, p. 6). Since the MOBILE model has many limitations with respect to I/M emissions reductions, one should be careful in applying it to evaluate an I/M program.

An intra-program evaluation, which compares initial and final I/M measurements under the same test procedures and conditions, seeks to reduce the impact of these variables. The drawback of the intra-program evaluation is that one cannot compare the emission reductions as measured in the program with the emission reductions, and associated SIP credits, as modeled by MOBILE.

The Sierra methodology assumes that a benchmark I/M program can be found that achieves the emissions reductions forecast in the MOBILE model. All other I/M programs would then be compared to the benchmark program, and SIP credit would be based on this comparison. However, if the emissions reductions from the benchmark program are overestimated, then the effect of using a benchmark is to "lower the bar" in terms of emissions reductions. There is evidence that even a seemingly well run centralized IM240 program, that meets EPA's recommendations for enhanced I/M, does not achieve the emissions reductions forecast in the MOBILE model (Harrington et al 1998).

An alternative method to evaluate I/M programs is to measure directly the emissions of the fleet based on individual vehicle's initial and final I/M test results. This approach quantifies actual emissions reductions as measured by the program, rather than as projected by a computer model. This approach also accounts for the unique nature of the vehicle fleet, fuel composition, and environmental conditions of any I/M area. In addition, the long-term benefit of the program can be readily measured by examining individual vehicle emissions test results over multiple I/M cycles. The drawback of this approach is that it gives no credit to cumulative emissions reductions from past I/M programs. It is thought that it will be easier for states with no previous I/M program to obtain substantial emissions reductions by adopting I/M, and that states that have had an I/M program operating for years will have little additional excess emissions available to reduce through an enhanced I/M program. However, there is evidence that there is little, if any, cumulative effect of older, basic I/M programs on emissions. For example, comparison of first-year average emissions by MY from different I/M programs implementing IM240 testing reveals that the program that had no prior I/M program (Ohio) has virtually the same average emissions as the program that previously was a basic I/M program (Wisconsin). If basic I/M had a cumulative effect on vehicle emissions, the Wisconsin program would have lower average emissions than the Ohio program. There is additional evidence that the cumulative emissions reductions from enhanced I/M programs is smaller than assumed in the MOBILE model. For instance, analysis of multiple years of IM240 data from two enhanced I/M states (Arizona and Colorado) indicates that as much as 40% of the vehicles that failed and were repaired in the first cycle of enhanced I/M failed in the second cycle (Wenzel 1998b, ENVIRON 1998).

Sierra's approach requires normalizing fleet emissions of the benchmark program to the MY-distribution of the evaluated fleet. However, the goal of I/M programs is to reduce overall emissions, not emissions by MY. The intra-state evaluation approach allows states to reduce fleet emissions by encouraging more rapid vehicle turnover (affecting MY distribution), and by reducing VMT. A more stringent I/M program will encourage a more rapid retirement of high emitting vehicles. EPA should recommend that states calculate fleet emissions reductions weighted by vehicles, as well as by VMT. States should be given credit for reducing overall fleet emissions, and not just reducing emissions by MY.

A Better Evaluation Methodology

We understand that EPA, and by association its contractor Sierra, are operating under the constraints of past decisions regarding how in-use emissions are modeled and how State Implementation Plan emission credits are quantified. However, we would like to summarize our recommendations on what we believe is a better method to evaluate a state's I/M program.

In addition to the intra-state evaluation of in-program I/M data, EPA should consider requiring states to use out-of-program data to evaluate their I/M program. Out-of-program data are needed to evaluate a program because they represent unscheduled testing of vehicle emissions. Remote sensing, if done carefully, is an efficient way to obtain in-use emissions measurements on a large fraction of the in-use fleet. Site selection is critical to using remote sensing data for program evaluation. Only use sites where vehicles are in controlled operating modes (or measure vehicle operating characteristics, such as speed, acceleration, and perhaps catalyst or engine temperature); utilize a small number of sites, to reduce site effects over long periods of time; and strategically locate sites to cover a wide range of socioeconomic levels. If done carefully, remote sensing offers several advantages over roadside emissions testing: if measurements are used for program evaluation only, and not for clean screening or high emitter identification, self-selection bias is reduced--participation is virtually mandatory, and there is no incentive to avoid the test. In addition, remote sensing is less expensive than roadside testing. Roadside testing offers some advantages over remote sensing. Visual inspections can be performed and the emissions test used can be the same as the I/M test.

States are already required to do random unscheduled testing of the I/M fleet, using either remote sensing or roadside pullover tailpipe testing, to monitor program effectiveness. However, this testing is required on only 0.5% of the I/M fleet (CFR 51.371). Because even carefully controlled remote sensing measurements are relatively inexpensive, measurements on thousands of vehicles, including multiple measurements of individual vehicles over time, can be made. Such data is necessary to perform a thorough evaluation of an I/M program. The optimal evaluation program would include data from all three types of measurement (in-program I/M testing, a large database of remote sensing data, and roadside pullovers for visual inspections).

An analysis of remote sensing data is likely to result in lower estimates of emission reductions than using in-program data. Remote sensing data are collected days, weeks, and months after vehicles pass out of the I/M program. Therefore it is likely that post-repair emissions measured by RSD will be lower than those measured by IM240, as will calculated emissions reductions due to the I/M program. Our analysis of 18 months of IM240 and remote sensing data in Arizona found that the CO emissions of the fleet entering the program were reduced by 15%, as measured by IM240 tests, and by 12% as measured by remote sensing. Emissions of vehicles failing their initial I/M test and getting repaired were reduced by 64% and 36% under each test type. However, the repair effectiveness diminishes as vehicles get further from their final I/M test. The reduction in CO emissions for the entire fleet, as measured by remote sensing, over one year after final I/M testing was only 6%, half of what it was immediately after final I/M testing (the reduction in emissions from repaired vehicles also was dramatically reduced, to 28%). Similarly, multiple years of Arizona in-program data show a similar reduction in repair effectiveness over time. 40% of the vehicles that fail in the first I/M cycle fail in the next I/M cycle, as opposed to 7% of the vehicles that pass the first I/M cycle (Wenzel 1998b). This reduction in repair effectiveness is either due to ineffective repairs, or different emissions controls breaking on the same vehicles. More research is needed to determine which of these factors is the major contributor. In any event, reduction in repair effectiveness over time is a real effect of the I/M program, and should be included in program evaluations.

On the other hand, an evaluation using remote sensing could <u>increase</u> calculated emissions reductions. Remote sensing data can also be collected for days, weeks, and months <u>prior to</u> the initial I/M test of individual vehicles. There is evidence that fleet average emissions increase as vehicles get closer to their initial I/M test. Much of this increase is due to emissions controls failing on a larger number of vehicles; some of this increase is due to aging and mileage accumulation of individual vehicles. In the weeks before the scheduled I/M test, however, average emissions are somewhat reduced. This is likely due to vehicle owners making repairs prior to their initial I/M test, to ensure that their vehicles pass the test, even in a centralized IM240 program such as that in Arizona. Our analysis of vehicles that fail their initial I/M test shows a reduction in average emissions immediately prior to the test. These reductions are not enough to cause the vehicles to pass the I/M test, however (Wenzel 1998b). Emissions reductions calculated on the basis of initial and final I/M test results would not include the emissions reductions from pre-test repairs. States should be given credit for emissions reductions due to repairs prior to the I/M test; these repairs likely would not have been made in the absence of an I/M program. The only way to measure the effect of these repairs is with unscheduled testing.

By focusing on the emissions of vehicles only after they pass through the I/M program, Sierra's proposed methodology ignores the effect of vehicles that never receive a passing I/M test on overall fleet emissions. Our analysis of remote sensing data in Arizona found that these vehicles account for about 3% of all vehicles, and 33% of all vehicles failing their initial I/M test, and about 14% of all CO emissions (Wenzel 1998b).

Out-of-program data can also be used to determine if vehicles that never receive a passing I/M test are still being driven in the I/M area. About 6% of the Arizona fleet measured by remote sensing prior to I/M testing fail their upcoming I/M test and pass subsequent retesting ("repaired" vehicles), while 3% of the fleet would fail their upcoming I/M test and never pass a subsequent test ("disappearing" vehicles). One year after final I/M testing, only 4% of the fleet observed by remote sensing were repaired vehicles (a 30% reduction), while only 1% of the fleet seen by remote sensing were disappearing vehicles (a 66% reduction; Wenzel 1998b). The disappearing vehicles appear to be dropping out of the Arizona I/M fleet at a faster rate than the repaired vehicles;

however, one-third of the disappearing vehicles were still being driven in the I/M area over one year after their I/M test. Out-of-program data are needed to identify the magnitude of this problem, and to identify vehicles that are circumventing the I/M program. It is not necessary to use remote sensing to accomplish this; video cameras that record license plates can be used as well (although remote sensing is necessary to estimate the emissions consequences of allowing these vehicles to continue operating in the I/M area).

References

Barth, Matthew, et al. *Development of a Modal Emissions Model: Final Report for NCHRP 25-11*. Prepared for the National Cooperative Highway Research Program of the Transportation Research Board. Draft.

ENVIRON. 1998. *Performance Audit of the Colorado AIR Program: Final Report.* Prepared for the State of Colorado, Office of State Auditor. March.

Harrington, Winston, Virginia McConnell, and Matthew Cannon. 1998. *A Behavioral Analysis of EPA's MOBILE Emission Factor Model*. Resources for the Future Discussion Paper 98-47, September.

Heirigs, Philip, and Jay Gordon. "Preconditioning Effects on I/M Test Results Using IM240 and ASM Procedures." Society of Automotive Engineers 960064.

Joy, Richard. 1998. "Improvements in Fast Pass, NOx Correction, and Reverse Dilution." Presentation at the 14th Annual Mobile Sources/Clean Air Conference, Breckenridge, Colorado, September 15-18.

Knepper, Jay, et al. 1992. "Fuel Effects in Auto/Oil High Emitting Vehicles." Presentation at the 3rd Annual CRC On-Road Vehicle Emissions Workshop, San Diego, California, December 1-3.

Stedman, Donald, et al. 1994. *On-Road Remote Sensing of CO and HC Emissions in California.* Final Report for the California Air Resources Board, Contract No. A032-093. Submitted by the University of Denver. February.

Wenzel, Tom and Marc Ross. 1997. "I/M Failure Rates by Vehicle Model." Presentation at the 7th CRC On-Road Vehicle Emissions Workshop, San Diego, California, April 9-11.

Wenzel, Tom, Robert Sawyer, and Marc Ross. 1998. "Seasonal Effects, Arizona IM240." Unpublished memo to US EPA Office of Mobile Sources, May 1.

Wenzel, Tom. 1998a. "Comparison of Emissions by Model using IM240 and Remote Sensing Data." Unpublished memo to US EPA Office of Mobile Sources, July 24.

Wenzel, Tom. 1998b. "Evaluation of Arizona's Enhanced I/M Program." Presentation at the 14th Annual Mobile Sources/Clean Air Conference, Breckenridge, Colorado, September 15-18.